Qualitative Representation of Dynamic Attributes of Trajectories

Tales Paiva Nogueira, Hervé Martin

Qualitative Representation of Dynamic Attributes of Trajectories

Tales Paiva Nogueira Hervé Martin
Univ. Grenoble Alpes, LIG, F-38000 Grenoble, France
CNRS, LIG, F-38000 Grenoble, France
paivan@imag.fr herve.martin@imag.fr

Abstract

Trajectory dynamic characteristics may be a very relevant source of information to analyze the behaviour of moving objects. However, most of existing works on trajectory representation deal only with basic parameters of trajectories, namely space and time. In this paper, we show how some derivatives of the spatio-temporal dimension, e.g. speed, acceleration, direction, may be integrated in trajectory modelling. We address the problem of representing trajectories in a way that qualitative descriptions of trajectories are stored and easily accessed through an ontology called QualiTraj which is also flexible enough to support relevant raw data representation. We validate our proposal with real GPS traces collected from a well-known sports tracking mobile application.
Keywords: geographical information systems, trajectory analysis, semantic trajectories, dynamic, user profile, ontologies

1 Introduction

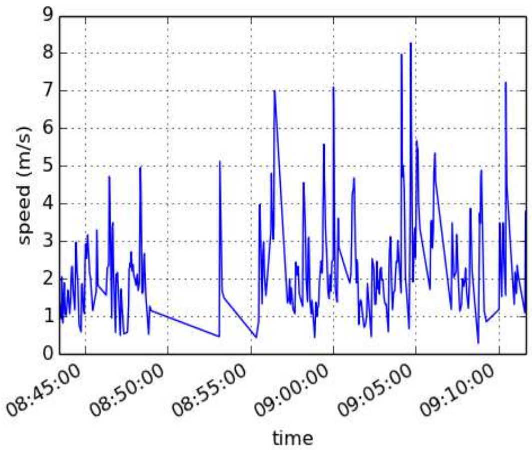
The Internet as we know it today is constantly evolving to a more and more connected system thanks to the increasing quantity of data made available as linked data, building what is called the Semantic Web [2]. With semantic web technologies we are moving from the web of documents towards the web of data, where machines will be able to understand and reason about the connections among different datasets and therefore enable the development of richer applications. It is of common knowledge that ontologies are well suited to represent datasets in this new paradigm.

At the same time, we are witnessing the development of mobile technologies such as smart phones for the acquisition of data in conjunction with many sensors as well as the grow of technologies of geolocation (GPS, A-GPS, GLONASS) and identification (RFID). The convergence of these technologies allows the easy acquisition of information about the trajectories of users using mobile devices. The acquisition, management, modelling, and analysis of such data provide many challenges related to the integration of these data with systems that already exist. Therefore, it should be taken into account that the multidimensional and multifaceted aspects of these data potentially holds a very rich semantics. There is a vast amount of works that propose to bridge the gap between trajectory representation and Semantic Web technologies, mainly regarding the representation with ontologies [1, 3, 9, 10, 13, 17, 19, 20].

The identification of mobility patterns has been an constant topic of interest in the GIScience area in several domains like tourism, road traffic, crisis management, marketing, etc. [8]. Several works have dealt with trajectory analysis proposing new ways of comparing, segmenting and clustering moving object's paths. But most of them only handle the geometric aspect of trajectories and just a few deal with dynamic parameters like speed and acceleration explicitly [4, 12].

In most cases, the variability of dynamic properties is very high. Take as example the raw representation of speed of a runner in Figure 1. Although it is obvious that the movement did not suffer exactly the same variations as it is depicted in the graphic due to the intrinsic error and noise of GPS readings, we can observe that there is a need to simplify this information so it can become more useful. The development of new methods and tools to analyse movement components like the one shown in Figure 1 is still a great challenge.

Figure 1: Time series of the speed profile of a runner captured by a smart phone application without any post processing treatment.



In this work, we argue that a qualitative representation of trajectories components are useful and enable new queries to be build and answer many application domain needs. The remainder of this paper is organized as follows: in section 2, we compare our work with similar proposals and highlight the differences between them. In section 3, we define what are the dynamic aspects considered in this paper. In section 4, the QualiTraj ontology is introduced, followed by a case study in section 5.

2 Related work

In [14], Rehrl et. al. proposed a method for semantic processing of GPS traces where information is extracted from raw data. Based on the assumption that the basic parameters to express motion in space and time are velocity and course, they defined six motion patterns with associated rules. The patterns are the following: *stand still* characterizes the absence of motion and is assumed when the velocity is less than 1 m/s; *steady motion* represents the periods when there is motion with constant velocity and is distinguished when velocity is greater than 1 m/s and acceleration lies between −0.3 m/s² and 0.3 m/s². *Positive acceleration* happens when the velocity increases and is greater than 1 m/s and acceleration is greater than 0.3 m/s²; *negative acceleration* is similar but acceleration should be less than −0.3 m/s². *Positive course change* is identified when there is a course change rate above 0.4 º/s, and *negative course change* is determined when this change is below −0.4 º/s.

While this categorization has as objective to improve the level of abstraction of motion data, the authors rely too heavily in thresholds to characterize speed, acceleration, and course changes. In an heterogeneous dataset, this approach does not seem adequate as these thresholds may vary depending on the mean of transportation. In our work, we preferred the usage of statistic measures whenever it was possible to avoid relying on thresholds that depends on the nature of the data being analyzed.

In [9], van Hage et. al. presented the Simple Event Model (SEM) and its application in the maritime domain. In their use case, events are automatically recognized from the Automatic Identification System (AIS) raw data and represented as SEM instances. From that, it was possible to characterize some types of ship behavior like *slowing down*, *speeding up*, and *anchored*. Three types of data were collected in the form of time series: location, speed, and course. Due to the large dimensions of the tracked ships and the fact that they do not accelerate nor change their courses quickly, their movement are very regular and much more easy to compress by a piecewise linear algorithm. In our paper, instead of AIS data with speed information already included, we have at first just GPS raw data from which we have to calculate the speed profile. Moreover, the nature of motion data is very different: runners instead of ships. Runners may have a much more irregular speed, acceleration, and changes in course direction when compared to ships. Besides, we take a qualitative approach towards the characterization of movement.

3 What is dynamic?

One of the most important features of spatio-temporal systems is the ability to trace the path that a moving object follows during some time. A trajectory can be defined as the user defined record of the evolution of the position (perceived as a point) of an object that is moving in space during a given time interval in order to achieve a given goal [15]. A research topic that is constantly studied in the trajectory analysis domain is related to the representation of these spatio-temporal paths. While the representation of trajectories with ontologies have already been subject of many studies, the dynamic aspects of trajectories are generally just mentioned as simple attributes or even not mentioned. Most works about trajectory analysis limits themselves to the geometric representations of trajectories as a static curve [5].

The dynamic properties that we talk about in this paper may have different names among the literature. Dodge et. al. [5], for instance, call them movement parameters and separate them in three groups: primitive parameters, primary derivatives, and secondary derivatives. Each group is further organized in spatial, temporal and spatio-temporal dimensions. The primitive parameters are the ones that has been the subject of most studies in GIS (position and time). The primary derivatives are distance, direction, spatial extent, duration, travel time, speed and velocity. The secondary derivatives are spatial distribution, change of direction, sinuosity, temporal distribution, change of duration, acceleration, and approaching rate. In this work, we are going to focus on the speed derivative, as we believe that this aspect of trajectories is crucial to the characterization of the behaviour of a moving object.

4 The QualiTraj Ontology

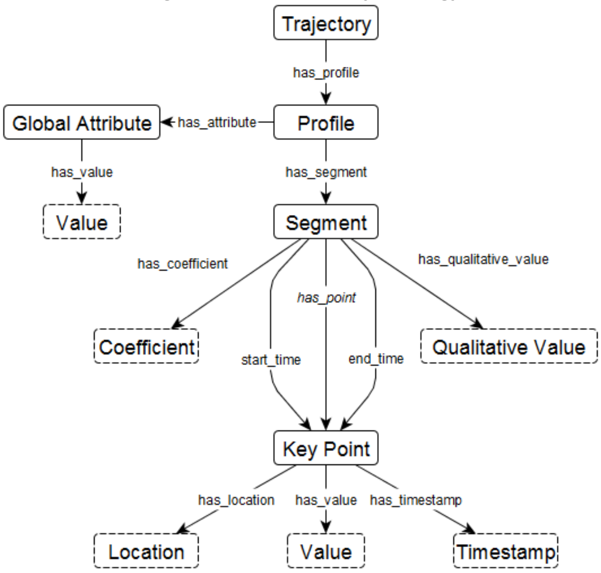
In this section, we present a modeling approach that enables the representation of trajectories' dynamic characteristics in a high abstraction level through an ontology. Figure 2 shows the basic structure of QualiTraj, the main contribution of this work. The top level element is the *Trajectory* entity, which represents a spatio-temporal path followed by a moving object. Each trajectory may have one or more profiles. Each *Profile* represents one dynamic aspect of a trajectory (e.g. speed, acceleration, direction).

Profiles may have aggregated measures that might be useful depending on the application. Thus, we included the *Global Attributes* entity to store information like the average speed of the whole trajectory.

The *Segment* is the entity that represents a relevant change in the dynamic property occurred along the trajectory. This element contains the qualitative information itself stored in the *Qualitative Value* entity. Each *Segment* starts and ends at a *Key Point*, i.e. a location in space and time that define the bounds of the segment. The *Key Points* may also be used to retrieve important information, e.g. where and when the highest speed was achieved. While this kind of data is not mandatory because it is application-specific and, on the other hand, the start and end points must always be represented, there are three relationships between *Segment* and *Key Point* in the ontology being optional only the one called *has_point*.

The kind of change is stored in the *Qualitative Value* entity associated with each *Segment*. The application developer should determine which values compose the lexical space of this entity. Another important element to represent each *Segment* is the *Coefficient*, an entity that holds the slope of the line that connects the starting and ending points. Having this information may be useful if we want to infer the approximate value of the profiled characteristic using a linear equation. The next section shows an example of the usage of the QualiTraj ontology in a real scenario.

Figure 2: The QualiTraj ontology



5 Case study

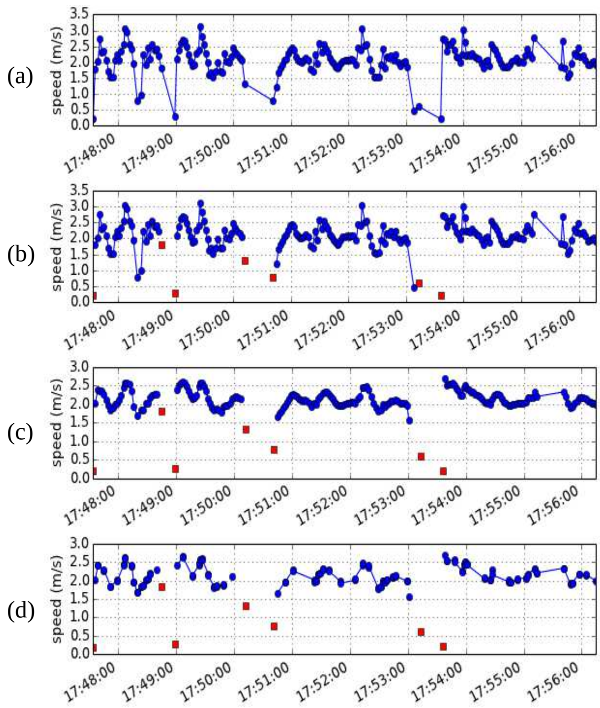
The studies about mobility analysis are numerous in the literature. In order to validate them, it is of vital importance to work with a representative dataset. The capture of real-life spatio-temporal data is generally expensive and time consuming due to the need of adequate equipment, search for subjects willing to participate (e.g. taxi drivers, shoppers, students), among other factors. Fortunately, there are some available datasets that can be freely downloaded, like GeoLife [21, 22], Reality Mining [6], geo-tagged photos from websites like Flickr[1] and Instagram[2], among others.

One interesting source of trajectory data is sports tracking websites and mobile applications, e.g. RunKeeper[3], Endomondo[4], Sports Tracker[5], Strava[6], MapMyRun[7], and have been source of studies like the one by Ferrari and Mamei [7]. Notwithstanding the widespread adoption of these services by professional athletes as well as by casual practitioners of sports activities, the collected data is not always publicly available. The minor part of sites provide an open API for third-party applications to access user-generated data.

For our case study, we collected data from users of MapMyRun application that shared their workouts publicly. We gathered information about 10 users that logged activities in the city of Grenoble, France, represented by 66 trajectories in total. In order to be clearer, we are going to show the raw data transformation steps of one short workout as it becomes easier to spot the changes suffered by the time series through all the steps. Figure 3 shows the raw speed data being pre-processed in order to simplify the stored data. The first graph shows the calculated speed at each point of the trajectory based on the latitude, longitude, and time difference between points. We have used the Haversine formula to calculate the approximate distance traveled during each sampled point. It is important to notice that a good speed approximation is heavily dependent on a good sampling rate, i.e. GPS fixes constantly recorded in small intervals of time.

The second step of the cleaning phase consists in detecting stops and moves of the tracked object. After that, we applied a Kalman filter [18] in order to smooth the data and attenuate GPS position errors. The last step of the smoothing phase is to summary the data points with a piecewise linear segmentation [11]. All the steps of the cleaning phase are depicted in Figure 3. In this specific example, the length of the time series was reduced from 60 points to only 28 points without losing the main characteristics of the signal.

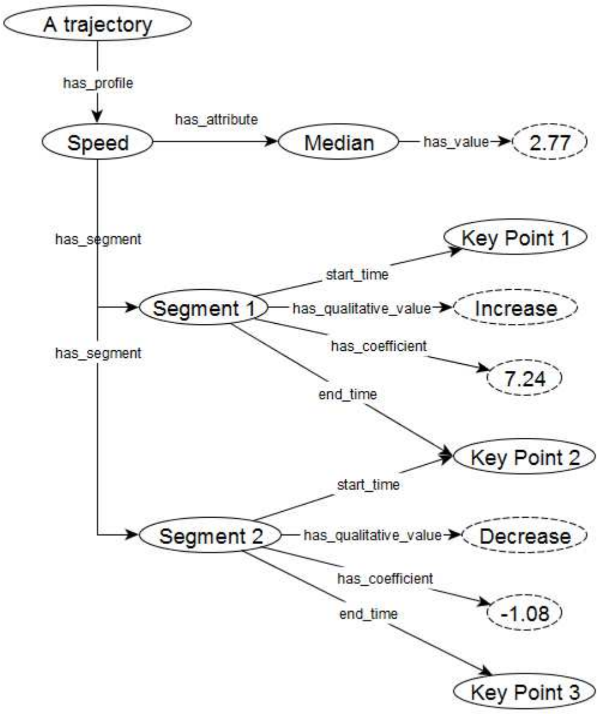
Figure 3: The evolution of speed during a four-minute walk and the steps of post-processing: (a) is the raw data, in (b) stops and moves are identified, (c) is the filtered signal, and (d) shows the piecewise linear segmentation result.



The final step in the speed representation of this trajectory consisted in creating the entities following the QualiTraj model. The lexical space used for the *Qualitative Value* entity was {"Increase", "Decrease", "Steady", and "Stop"}. Figure 4 shows the first two Segments represented with QualiTraj. The first segment consists in an increase of speed from 2.02 m/s to 2.40 m/s, which are the points of a line with an angle of 7.24 degrees. We omitted the timestamp and location of *Key* Points to improve the readability of the example. Notice that the same *Key Point*, "Key Point 2", has been reused in both segments, avoiding data duplication thanks to the graph structure of the ontological modeling approach.

[1] www.flickr.com
[2] www.instagram.com
[3] www.runkeeper.com
[4] www.endomondo.com
[5] www.sports-tracker.com
[6] www.strava.com
[7] www.mapmyrun.com

Figure 4: Part of qualitative representation of a trajectory using the QualiTraj ontology



6 Conclusion

In this paper, we demonstrated how it is possible to enrich raw trajectory data with dynamic aspects of movement and provide an infrastructure for querying this new knowledge through an ontology.

The representation of spatio-temporal data by means of ontologies is even more useful when the inference features of reasoners are explored. As a following activity of this work, we will investigate how reasoners can improve the analysis of patterns of dynamic movement parameters of trajectories. Queries that involve more than one moving object form and important group of queries about relative motion and should be studied in the future.

Another important development will be the connection of the proposed ontology with different linked data sources like the LinkedGeoData project [16], which provides OpenStreetMaps information in the format suitable for the Semantic Web. For instance, we could formulate queries in the domain of traffic analysis to find drivers that do not slow down near traffic-calming features as the OpenStreetMaps dataset has a traffic-calming key for many possible features of this type, like bumpers, chicanes, cushions, and others. In this way, more complex queries and reasoning tasks can be also envisaged as future work.

Acknowledgments

The authors would like to thank the French Ministry of Higher Education and Research (Ministère de l'Enseignement Supérieur et de la Recherche de la France – MESR) for supporting this work.

References

[1] Baglioni, M., Macedo, J., Renso, C. and Wachowicz, M. 2008. An ontology-based approach for the semantic modelling and reasoning on trajectories. *Advances in Conceptual Modeling – Challenges and Opportunities*. Springer Berlin Heidelberg. 344–353.

[2] Bizer, C., Heath, T. and Berners-Lee, T. 2009. Linked Data - The Story So Far. *International Journal on Semantic Web and Information Systems*. 5, 3 (Jan. 2009), 1–22.

[3] Camossi, E., Villa, P. and Mazzola, L. 2013. Semantic-based Anomalous Pattern Discovery in Moving Object Trajectories. *CoRR*. abs/1305.1, (May 2013).

[4] Dodge, S., Weibel, R. and Forootan, E. 2009. Revealing the physics of movement: Comparing the similarity of movement characteristics of different types of moving objects. *Computers, Environment and Urban Systems*. 33, 6 (Nov. 2009), 419–434.

[5] Dodge, S., Weibel, R. and Lautenschütz, A.-K. 2008. Towards a Taxonomy of Movement Patterns. (2008), 1–12.

[6] Eagle, N. and (Sandy) Pentland, A. 2005. Reality mining: sensing complex social systems. *Personal and Ubiquitous Computing*. 10, 4 (Nov. 2005), 255–268.

[7] Ferrari, L. and Mamei, M. 2013. Identifying and understanding urban sport areas using Nokia Sports Tracker. *Pervasive and Mobile Computing*. 9, 5 (Oct. 2013), 616–628.

[8] Güting, R.H., Böhlen, M.H., Erwig, M., Jensen, C.S., Lorentzos, N.A., Schneider, M. and Vazirgiannis, M. 2000. A foundation for representing and querying moving objects. *ACM Transactions on Database Systems*. 25, 1 (Mar. 2000), 1–42.

[9] Van Hage, W.R., Malaisé, V., de Vries, G., Schreiber, G. and van Someren, M. 2009. Combining ship trajectories and semantics with the simple event model (SEM). *Proceedings of the 1st ACM international workshop on Events in multimedia - EiMM '09* (New York, New York, USA, 2009), 73.

[10] Hu, Y., Janowicz, K., Carral, D., Scheider, S., Kuhn, W., Berg-Cross, G., Hitzler, P., Dean, M. and Kolas, D. 2013. A Geo-ontology Design Pattern for Semantic Trajectories. *Spatial Information Theory*. T. Tenbrink, J. Stell, A. Galton, and Z. Wood, eds. Springer International Publishing. 438–456.

[11] Keogh, E., Chu, S., Hart, D. and Pazzani, M. 2001. An online algorithm for segmenting time series.

Proceedings 2001 IEEE International Conference on Data Mining (2001), 289–296.

[12] Laube, P., Dennis, T., Forer, P. and Walker, M. 2007. Movement beyond the snapshot - dynamic analysis of geospatial lifelines. *Computers, Environment and Urban Systems*. 31, 5 (2007), 481–501.

[13] Parent, C., Pelekis, N., Theodoridis, Y., Yan, Z., Spaccapietra, S., Renso, C., Andrienko, G., Andrienko, N., Bogorny, V., Damiani, M.L., Gkoulalas-Divanis, A. and Macedo, J. 2013. Semantic trajectories modeling and analysis. *ACM Computing Surveys*. 45, 4 (Aug. 2013), 1–32.

[14] Rehrl, K., Leitinger, S., Krampe, S. and Stumptner, R. 2010. An Approach to Semantic Processing of GPS Traces. *Proceedings of the 1st Workshop on Movement Pattern Analysis* (Zurich, Switzerland, 2010), 136–142.

[15] Spaccapietra, S., Parent, C., Damiani, M.L., Macedo, J.A.F., Porto, F., Vangenot, C. and Demacedo, J. 2008. A conceptual view on trajectories. *Data & Knowledge Engineering*. 65, 1 (Apr. 2008), 126–146.

[16] Stadler, C., Lehmann, J., Höffner, K. and Auer, S. 2012. LinkedGeoData: A Core for a Web of Spatial Open Data. *Semantic Web journal*. 3, 4 (2012), 333–354.

[17] Wannous, R., Malki, J., Bouju, A. and Vincent, C. 2013. Modelling Mobile Object Activities Based on Trajectory Ontology Rules Considering Spatial Relationship Rules. *Modeling Approaches and Algorithms for Advanced Computer Applications*. A. Amine, A.M. Otmane, and L. Bellatreche, eds. Springer International Publishing. 249–258.

[18] Welch, G. and Bishop, G. 1995. *An introduction to the Kalman filter*.

[19] Yan, Z., Chakraborty, D., Parent, C., Spaccapietra, S. and Aberer, K. 2011. SeMiTri: A Framework for Semantic Annotation of Heterogeneous Trajectories. *Proceedings of the 14th International Conference on Extending Database Technology - EDBT/ICDT '11* (New York, New York, USA, 2011), 259.

[20] Yan, Z., Macedo, J., Parent, C. and Spaccapietra, S. 2008. Trajectory Ontologies and Queries. *Transactions in GIS*. 12, (Dec. 2008), 75–91.

[21] Zheng, Y., Xie, X. and Ma, W.-Y. 2010. GeoLife: A collaborative social networking service among user, location and trajectory. *IEEE Data Engineering Bulletin*. 49 (2010), 1–8.

[22] Zheng, Y., Zhang, L., Xie, X. and Ma, W.-Y. 2009. Mining interesting locations and travel sequences from GPS trajectories. *Proceedings of the 18th international conference on World wide web WWW 09*. 19, (2009), 791.